Kristianstad University
www.hkr.se

Fredrik Frisk
Computer Science Department

Content and purpose of this lab

The main focus of the lab is to do more classification.

Save the code you are writing in this lab for future use. To pass the lab you need to solve/program the different bullet points and be able to explain your results. If you are not finished with the all the bullet points the remaining ones are a part of the required preparations for part 4 of the labs. The lab report in the end is an individual report, but you are allowed to work two and two with one exception all of you have to record your own sensor data.

Required Preparations

- You need to finish part two of the lab and be able to show the result in the beginning of the lab.
- Install keras and tensorflow if you have not done it before. Create a new environment for this installation. We do not want to crash the old one. Keras and tensorflow is not needed for the preparation for the lab, but it is needed for the lab itself.

We will soon start with the neural network so som more training data is always good to have access to.

- Record every class one more time. That is for each of the seven (7) classes you should record an additional 15 s of data.
- Update the training folder so you include the new training data in each pickle-file. That is in the training folder you should have 7 pickle files one for each class. Each file contains at least 4500 instances with seven colums.

Be prepared to show/load the newly created pickle-files in the beginning of the lab.

- Use the gridsearchCV with KNN and DecisionTrees to classify the four movements classes. Work through as you did in lab part 2. That is in the end present the accuracy and confusion matrix for the optimal estimator applied on the test set.

Normalization

- See lab part 2. Normalise the attributes using min-max-scaler. Redo the classification above with the four movement classes. Compare the results with the previous results you got from the preparations.

Correlation

One can use correlation to investigate the dependence between variables/features.

- Calculate the correlation matrix between all features. From this result, discuss if all features are useful for the classification.

  We talk here about the four classes regarding movement and the six features we have worked with so far

Feature Transformation

- **Tip: store the transformed dataframes in new pickle files. You never know if you need them later.**

One way of getting better results is to make a feature transformation. We will use the standard deviation this time. You should use the rolling method in Pandas

https://pandas.pydata.org/docs/reference/api/pandas.core.window.rolling.Rolling.std.html

- Transform all features using rolling with the parameter 10. Tip: store the transformed dataframes in new pickle files. You never know if you need them later.
- What does the parameter (.rolling(10)) mean?
- Make the classification again (same classes as above) as previously, with KNN or Decision Trees. Compare the results with previous, untransformed features. Is the result improved?
- Try with a larger parameter for example 50. Is the result improved?

Neural Networks – At Last!

Now we will use the original features again. **Do not use the transformed features!**

- Build a feedforward neural network that has 6 features as input and 4 classes as output. The network should have 2 hidden layers. You should try the network with **10 and 50** neurons in each layer. That is, you need two different models.

As in the examples in the book divide the training set into a pure training set and a validation set. Leave the test set for prediction later

- Plot the accuracy and loss for both the training and validation set for both models, and decide which model is the best.

It is now time to work with the test set.

- Use the method evaluate on the test set and compare the results with the training set and validation set.

The following link explains the difference between evaluate and predict methods in keras

https://saturncloud.io/blog/keras-understanding-the-difference-between-modelevaluate-and-modelpredict-accuracy-in-multiclass-nlp-tasks/

For calculating the confusion matrix should you use evaluate or predict?

- Calculate the confusion matrix and plot it. Explain the result.